

OIPE

RAW SEQUENCE LISTING

DATE: 09/27/2001

PATENT APPLICATION: US/09/884,987

TIME: 14:13:38

Input Set : A:\0020-4877P.ST25.txt

Output Set: N:\CRF3\09272001\I884987.raw

3 <110> APPLICANT: NAGATA, Shigekazu et al
 5 <120> TITLE OF INVENTION: DNA CODING FOR HUMAN CELL SURFACE ANTIGEN
 7 <130> FILE REFERENCE: 0020-4877P
 9 <140> CURRENT APPLICATION NUMBER: US 09/884,987
 10 <141> CURRENT FILING DATE: 2001-06-21
 12 <160> NUMBER OF SEQ ID NOS: 11
 14 <170> SOFTWARE: PatentIn version 3.0
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 2534

ENTERED

18 <212> TYPE: DNA
 19 <213> ORGANISM: Homo sapiens
 21 <220> FEATURE:
 22 <221> NAME/KEY: polyA_site
 23 <222> LOCATION: (1831)..(1836)
 25 <220> FEATURE:
 26 <221> NAME/KEY: mat_peptide
 27 <222> LOCATION: (243)..()
 29 <220> FEATURE:
 30 <221> NAME/KEY: sig_peptide
 31 <222> LOCATION: (195)..(242)
 33 <220> FEATURE:
 34 <221> NAME/KEY: CDS
 35 <222> LOCATION: (195)..(1199)
 37 <220> FEATURE:
 38 <221> NAME/KEY: polyA_site
 39 <222> LOCATION: (2352)..(2357)
 41 <220> FEATURE:
 42 <221> NAME/KEY: polyA_site
 43 <222> LOCATION: (2518)..(2532)
 45 <400> SEQUENCE: 1

46	gacgcttctg gggagtgagg gaagcgggtt acgagtgact tggtcggagc ctcaggggag	60
48	ggcactggca cggaaacacac cctgaggcca gccctggctg cccaggcgga gctgcctctt	120
50	ctcccgcggg ttggtggacc cgtcagtagc ggagttgggg aagctctttc acttcggagg	180
52	attgctcaac aacc atg ctg ggc atc tgg acc ctc cta cct ctg gtt ctt	230
53	Met Leu Gly Ile Trp Thr Leu Leu Pro Leu Val Leu	
54	-15 -10 -5	
56	acg tct gtt gct aga tta tcg tcc aaa agt gtt aat gcc caa gtg act	278
57	Thr Ser Val Ala Arg Leu Ser Ser Lys Ser Val Asn Ala Gln Val Thr	
58	-1 1 5 10	
60	gac atc aac tcc aag gga ttg gaa ttg agg aag act gtt act aca gtt	326
61	Asp Ile Asn Ser Lys Gly Leu Glu Leu Arg Lys Thr Val Thr Thr Val	
62	15 20 25	
64	gag act cag aac ttg gaa ggc ctg cat cat gat ggc caa ttc tgc cat	374
65	Glu Thr Gln Asn Leu Glu Gly Leu His His Asp Gly Gln Phe Cys His	
66	30 35 40	
68	aag ccc tgt cct cca ggt gaa agg aaa gct agg gac tgc aca gtc aat	422
69	Lys Pro Cys Pro Pro Gly Glu Arg Lys Ala Arg Asp Cys Thr Val Asn	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,987

DATE: 09/27/2001

TIME: 14:13:38

Input Set : A:\0020-4877P.ST25.txt

Output Set: N:\CRF3\09272001\I884987.raw

70	45		50		55		60	
72	ggg	gat	gaa	cca	gac	tgc	gtg	ccc
73	Gly	Asp	Glu	Pro	Asp	Cys	Val	Pro
74				65			70	
76	gac	aaa	gcc	cat	ttt	tct	tcc	aaa
77	Asp	Lys	Ala	His	Phe	Ser	Ser	Lys
78				80			85	
80	gaa	gga	cat	ggc	tta	gaa	gtg	gaa
81	Glu	Gly	His	Gly	Leu	Glu	Val	Glu
82				95			100	
84	acc	aag	tgc	aga	tgt	aaa	cca	aac
85	Thr	Lys	Cys	Arg	Cys	Lys	Pro	Asn
86				110			115	
88	gaa	cac	tgt	gac	cct	tgc	acc	aaa
89	Glu	His	Cys	Asp	Pro	Cys	Thr	Lys
90	125						130	
92	tgc	aca	ctc	acc	agc	aac	acc	aag
93	Cys	Thr	Leu	Thr	Ser	Asn	Thr	Lys
94							145	
96	aac	ttg	ggg	tgg	ctt	tgt	ctt	ctt
97	Asn	Leu	Gly	Trp	Leu	Cys	Leu	Leu
98				160			165	
100	tgg	gtg	aag	aga	aag	gaa	gta	cag
101	Trp	Val	Lys	Arg	Lys	Glu	Val	Gln
102				175			180	
104	gaa	aac	caa	ggt	tct	cat	gaa	tct
105	Glu	Asn	Gln	Gly	Ser	His	Glu	Ser
106				190			195	
108	gca	ata	aat	tta	tct	gat	gtt	gac
109	Ala	Ile	Asn	Leu	Ser	Asp	Val	Asp
110	205						210	
112	gct	gga	gtc	atg	aca	cta	agt	caa
113	Ala	Gly	Val	Met	Thr	Leu	Ser	Gln
114							225	
116	ggt	gtc	aat	gaa	gcc	aaa	ata	gat
117	Gly	Val	Asn	Glu	Ala	Lys	Ile	Asp
118				240			245	
120	gac	aca	gca	gaa	cag	aaa	gtt	caa
121	Asp	Thr	Ala	Glu	Gln	Lys	Val	Gln
122				255			260	
124	cat	gga	aag	aaa	gaa	gcg	tat	gac
125	His	Gly	Lys	Lys	Glu	Ala	Tyr	Asp
126				270			275	
128	gcc	aat	ctt	tgt	act	ctt	gca	gag
129	Ala	Asn	Leu	Cys	Thr	Leu	Ala	Glu
130	285						290	
132	gac	att	act	agt	gac	tca	gaa	aat
133	Asp	Ile	Thr	Ser	Asp	Ser	Glu	Asn
134							305	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,987

DATE: 09/27/2001

TIME: 14:13:38

Input Set : A:\0020-4877P.ST25.txt

Output Set: N:\CRF3\09272001\I884987.raw

```

136 agc ttg gtc tagagtga aaacaacaaat tcagttctga gtatatgcaa      1239
137 Ser Leu Val
140 ttagtggttg aaaagattct taatagctgg ctgtaaatac tgcttggtt tttactgggt      1299
142 acattttatc atttattagc gctgaagagc caacatatct gtagattttt aatatctcat      1359
144 gattctgect ccaaggatgt ttaaaatcta gttgggaaaa caaacttcat caagagtaaa      1419
146 tgcagtggca tgctaagtac ccaaatagga gtgtatgcag aggatgaaag attaaagatta      1479
148 tgctctggca tctaacatat gattctgtag tatgaatgta atcagtgtat gttagtacaa      1539
150 atgtctatcc acaggctaac cccactctat gaatcaatag aagaagctat gaccttttgc      1599
152 tgaaatatca gttactgaac aggcaggcca ctttgctctt aaattacctc tgataattct      1659
154 agagatttta ccatatttct aaactttgtt tataactctg agaagatcat atttatgtaa      1719
156 agtatatgta tttgagtga gaatttaaat aaggtctctac ctcaaagacc ttgcacagt      1779
158 ttaltgggtg catattatac aatatttcaa ttgtgaattc acatagaaaa cattaaatta      1839
160 taatgtttga ctattatata tgtgtatgca ttttactggc tcaaaactac ctacttcttt      1899
162 ctgaggcatc aaaagcattt tgagcaggag agtattacta gagctttgac acctctccat      1959
164 tttlgccttg gtgctcatct taatggccta atgcaccccc aaacatggaa atatcaccaa      2019
166 aaaatactta atagtccacc aaaaggcaag actgccccta gaaattctag cctggtttgg      2079
168 agatactaac tgctctcaga gaaagttagt ttgtgacatg tcatgaacct atgtttgcaa      2139
170 tcaaagatga taaaatagat ttttattttt cccccacccc cgaaaatggt caataatgtc      2199
172 ccatgtaaaa cctgctacaa atggcagctt atacatagca atggtaaaaat catcatctgg      2259
174 atttaggaat tgctcttgtc ataccctcaa gtttctaaga tttaagattc tccttactac      2319
176 taccctacgt ttaaatatct ttgaaagttt gtattaaatg tgaattttta gaaataatat      2379
178 ttatatttct gtaaatgtaa actgtgaaga tagttataaa ctgaagcaga tacctggaac      2439
180 cacctaaaga acttccattt atggaggatt tttttgcccc ttgtggttgg aattataaaa      2499
182 tataggtaaa agtacgtaat taaataatgt ttttg      2534
185 <210> SEQ ID NO: 2
186 <211> LENGTH: 335
187 <212> TYPE: PRT
188 <213> ORGANISM: Homo sapiens
190 <400> SEQUENCE: 2
192 Met Leu Gly Ile Trp Thr Leu Leu Pro Leu Val Leu Thr Ser Val Ala
193 -15 -10 -5 -1
196 Arg Leu Ser Ser Lys Ser Val Asn Ala Gln Val Thr Asp Ile Asn Ser
197 1 5 10 15
200 Lys Gly Leu Glu Leu Arg Lys Thr Val Thr Thr Val Glu Thr Gln Asn
201 20 25 30
204 Leu Glu Gly Leu His His Asp Gly Gln Phe Cys His Lys Pro Cys Pro
205 35 40 45
208 Pro Gly Glu Arg Lys Ala Arg Asp Cys Thr Val Asn Gly Asp Glu Pro
209 50 55 60
212 Asp Cys Val Pro Cys Gln Glu Gly Lys Glu Tyr Thr Asp Lys Ala His
213 65 70 75 80
216 Phe Ser Ser Lys Cys Arg Arg Cys Arg Leu Cys Asp Glu Gly His Gly
217 85 90 95
220 Leu Glu Val Glu Ile Asn Cys Thr Arg Thr Gln Asn Thr Lys Cys Arg
221 100 105 110
224 Cys Lys Pro Asn Phe Phe Cys Asn Ser Thr Val Cys Glu His Cys Asp
225 115 120 125
228 Pro Cys Thr Lys Cys Glu His Gly Ile Ile Lys Glu Cys Thr Leu Thr
229 130 135 140

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,987

DATE: 09/27/2001

TIME: 14:13:38

Input Set : A:\0020-4877P.ST25.txt

Output Set: N:\CRF3\09272001\I884987.raw

```

232 Ser Asn Thr Lys Cys Lys Glu Glu Gly Ser Arg Ser Asn Leu Gly Trp
233 145 150 155 160
236 Leu Cys Leu Leu Leu Pro Ile Pro Leu Ile Val Trp Val Lys Arg
237 165 170 175
240 Lys Glu Val Gln Lys Thr Cys Arg Lys His Arg Lys Glu Asn Gln Gly
241 180 185 190
244 Ser His Glu Ser Pro Thr Leu Asn Pro Glu Thr Val Ala Ile Asn Leu
245 195 200 205
248 Ser Asp Val Asp Leu Ser Lys Tyr Ile Thr Thr Ile Ala Gly Val Met
249 210 215 220
252 Thr Leu Ser Gln Val Lys Gly Phe Val Arg Lys Asn Gly Val Asn Glu
253 225 230 235 240
256 Ala Lys Ile Asp Glu Ile Lys Asn Asp Asn Val Gln Asp Thr Ala Glu
257 245 250 255
260 Gln Lys Val Gln Leu Leu Arg Asn Trp His Gln Leu His Gly Lys Lys
261 260 265 270
264 Glu Ala Tyr Asp Thr Leu Ile Lys Asp Leu Lys Lys Ala Asn Leu Cys
265 275 280 285
268 Thr Leu Ala Glu Lys Ile Gln Thr Ile Ile Leu Lys Asp Ile Thr Ser
269 290 295 300
272 Asp Ser Glu Asn Ser Asn Phe Arg Asn Glu Ile Gln Ser Leu Val
273 305 310 315
276 <210> SEQ ID NO: 3
277 <211> LENGTH: 119
278 <212> TYPE: PRT
279 <213> ORGANISM: Homo sapiens
281 <400> SEQUENCE: 3
283 Gln Asn Leu Glu Gly Leu His His Asp Gly Gln Phe Cys His Lys Pro
284 1 5 10 15
286 Cys Pro Pro Gly Glu Arg Lys Ala Arg Asp Cys Thr Val Asn Gly Asp
287 20 25 30
289 Glu Pro Asp Cys Val Pro Cys Gln Glu Gly Lys Glu Tyr Thr Asp Lys
290 35 40 45
292 Ala His Phe Ser Ser Lys Cys Arg Arg Cys Arg Leu Cys Asp Glu Gly
293 50 55 60
295 His Gly Leu Glu Val Glu Ile Asn Cys Thr Arg Thr Gln Asn Thr Lys
296 65 70 75 80
298 Cys Arg Cys Lys Pro Asn Phe Phe Cys Asn Ser Thr Val Cys Glu His
299 85 90 95
301 Cys Asp Pro Cys Thr Lys Cys Glu His Gly Ile Ile Lys Glu Cys Thr
302 100 105 110
304 Leu Thr Ser Asn Thr Lys Cys
305 115
307 <210> SEQ ID NO: 4
308 <211> LENGTH: 153
309 <212> TYPE: PRT
310 <213> ORGANISM: Homo sapiens
312 <400> SEQUENCE: 4
314 Val Cys Pro Gln Gly Lys Tyr Ile His Pro Gln Asn Asn Ser Ile Cys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,987

DATE: 09/27/2001

TIME: 14:13:38

Input Set : A:\0020-4877P.ST25.txt

Output Set: N:\CRF3\09272001\I884987.raw

```

315 1           5           10           15
317 Cys Thr Lys Cys His Lys Gly Thr Tyr Leu Tyr Asn Asp Cys Pro Gly
318           20           25           30
320 Pro Gly Gln Asp Thr Asp Cys Arg Glu Cys Glu Ser Gly Ser Phe Thr
321           35           40           45
323 Ala Ser Glu Asn His Leu Arg His Cys Leu Ser Cys Ser Lys Cys Arg
324           50           55           60
326 Lys Glu Met Gly Gln Val Glu Ile Ser Ser Cys Thr Val Asp Arg Asp
327 65           70           75           80
329 Thr Val Cys Gly Cys Arg Lys Asn Gln Tyr Arg His Tyr Trp Ser Glu
330           85           90           95
332 Asn Leu Phe Gln Cys Phe Asn Cys Ser Leu Cys Leu Asn Gly Thr Val
333           100          105          110
335 His Leu Ser Cys Gln Glu Lys Gln Asn Thr Val Cys Thr Cys His Ala
336           115          120          125
338 Gly Phe Phe Leu Arg Glu Asn Glu Cys Val Ser Cys Ser Asn Cys Lys
339           130          135          140
341 Lys Ser Leu Glu Cys Thr Lys Leu Cys
342 145          150
344 <210> SEQ ID NO: 5
345 <211> LENGTH: 163
346 <212> TYPE: PRT
347 <213> ORGANISM: Homo sapiens
349 <400> SEQUENCE: 5
351 Thr Cys Arg Leu Arg Glu Tyr Tyr Asp Gln Thr Ala Gln Met Cys Cys
352 1           5           10           15
354 Ser Lys Cys Ser Pro Gly Gln His Ala Lys Val Phe Cys Thr Lys Thr
355           20           25           30
357 Ser Asp Thr Val Cys Asp Ser Cys Glu Asp Ser Thr Tyr Thr Gln Leu
358           35           40           45
360 Trp Asn Trp Val Pro Glu Cys Leu Ser Cys Gly Ser Arg Cys Ser Asp
361           50           55           60
363 Asp Gln Val Glu Thr Gln Ala Cys Thr Arg Glu Gln Asn Arg Ile Cys
364 65           70           75           80
366 Thr Cys Arg Pro Gly Trp Tyr Cys Ala Leu Ser Lys Gln Glu Gly Cys
367           85           90           95
369 Arg Leu Cys Ala Pro Leu Arg Lys Cys Arg Pro Gly Phe Gly Val Ala
370           100          105          110
372 Arg Pro Gly Thr Glu Thr Ser Asp Val Val Cys Lys Pro Cys Ala Pro
373           115          120          125
375 Gly Thr Phe Ser Asn Thr Thr Ser Ser Thr Asp Ile Cys Arg Pro His
376           130          135          140
378 Gln Ile Cys Asn Val Val Ala Ile Pro Gly Asn Ala Ser Met Asp Ala
379 145          150          155          160
381 Val Cys Thr
384 <210> SEQ ID NO: 6
385 <211> LENGTH: 159
386 <212> TYPE: PRT
387 <213> ORGANISM: Homo sapiens

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/884,987

DATE: 09/27/2001

TIME: 14:13:39

Input Set : A:\0020-4877P.ST25.txt

Output Set: N:\CRF3\09272001\I884987.raw